

Claims

We claim:

1. A method for developing a data model in a data mining system, comprising:
 - providing a database table of predefined data transformations;
 - providing raw data;
 - developing a data model of variables using at least one data transformation selected from the database table and the raw data; and
 - writing a specification for applying the data model operationally.
2. The method of claim 1, further comprising coding and deploying the data model using the specification.
3. The method of claim 1, wherein the developing step comprises:
 - determining a set of variables for a desired prediction, wherein the set of variables include at least one predefined data transformation selected from the database table; and
 - deriving a mathematical relationship between the set of variables.
4. The method of claim 3, wherein the developing step further comprises applying the raw data to the set of variables.

5. The method of claim 1, wherein the database table of predefined data transformations associates each of the predefined data transformations with a unique identifier, a description and a validity period.
6. The method of claim 5, wherein the developing step comprises retrieving the at least one predefined data transformation from the database table according to its unique identifier.
7. The method of claim 1, wherein the step of providing raw data comprises extracting raw data from a data warehouse.
8. The method of claim 1, wherein the step of writing a specification comprises writing a reusable set of instructions for applying the data model operationally.
9. The method of claim 1, further comprising providing a modification policy that governs modification of the predefined data transformations in the database table.

10. A computer-implemented method for developing a data model in a data mining system, comprising:

providing a database table of predefined data transformations, wherein each of the predefined data transformations is associated in the database table with a unique identifier and a description;

extracting raw data from a data warehouse;

determining a set of variables for a desired prediction, wherein the set of variables comprise at least one predefined data transformation selected from the table;

developing a data model for the desired prediction by applying the raw data to the set of variables and deriving a mathematical relationship between the set of variables; and

writing a specification for applying the data model operationally.

11. The method of claim 10, further comprising coding and deploying the data model using the specification.

12. The method of claim 10, wherein each of the predefined data transformation is further associated with a validity period.

13. The method of claim 10, wherein the determining step comprises retrieving the at least one predefined data transformation from the database table according to its unique identifier.

14. The method of claim 10, wherein the step of writing a specification comprises writing a set of instructions for applying the set of variables of the data model operationally.

15. The method of claim 10, wherein the specification is reusable

16. The method of claim 10, further comprising providing a modification policy that governs modification of the predefined data transformations in the database table.

17. A database table for developing a data model in a data mining system comprising a set of entries, wherein each of the set of entries includes a predefined data transformation, a unique identifier for the predefined data transformation, a description of the predefined data transformation and a validity period for the predefined data transformation.

18. The database table of claim 17, wherein the predefined data transformation comprises a mathematical expression for determining a result.

19. The database table of claim 17, wherein the predefined data transformation comprises a SQL expression.

20. The database table of claim 17, wherein the validity period comprises a start date and a termination date for the predefined data transformation.

21. The database table of claim 17, wherein the description comprises meta data.

22. The database table of claim 17, wherein the database table is associated with a modification policy that governs modification of the predefined data transformation.

23. A computerized system for developing a data model in a data mining system, comprising:
- a database table of predefined data transformations;
 - a data import system for extracting raw data from a data warehouse;
 - a variable determination system for determining a set of variables for a desired prediction,
- wherein the set of variables comprises at least one predefined data transformation selected from the database table;
- a model development system for developing a data model for the desired prediction using the determined variables; and
 - a specification development system for developing a specification for applying the data model operationally.
24. The system of claim 23, further comprising an output system for outputting the data model for coding and deployment based on the specification.
25. The system of claim 23, wherein the model development system applies the raw data to the set of variables and derives a mathematical relationship between the set of variables.
26. The system of claim 23, wherein the database table of predefined data transformations associates each of the predefined data transformations with a unique identifier, a description and a validity period.

27. The system of claim 26, wherein the variable determination system retrieves the at least one predefined data transformation from the database table according to its unique identifier.

28. The system of claim 23, wherein the specification comprises a reusable set of instructions for applying the variables of the data model operationally.

29. The system of claim 23, further comprising a modification policy that governs modification of the predefined data transformations in the database table.

30. A program product stored on a recordable medium for developing a data model in a data mining system, which when executed, comprises:

program code for extracting raw data from a data warehouse;

program code for determining a set of variables for a desired prediction, wherein the set of variables comprises at least one predefined data transformation selected from a database table of predefined data transformations;

program code for developing a data model for the desired prediction using the determined variables; and

program code for developing a specification for applying the data model operationally.

31. The program product of claim 30, further comprising program code for outputting the data model for coding and deployment based on the specification.

32. The program product of claim 30, wherein the program code for developing a data model applies the raw data to the set of variables and derives a mathematical relationship between the set of variables.

33. The program product of claim 30, wherein the database table of predefined data transformations associates each of the predefined data transformations with a unique identifier, a description and a validity period.

34. The program product of claim 33, wherein the program code for determining retrieves the at least one predefined data transformation from the database table according to its unique identifier.

35. The program product of claim 30, wherein the specification comprises a re-useable set of instructions for applying the variables of the data model operationally.

36. The program product of claim 30, further comprising a modification policy that governs modification of the predefined data transformations in the database table.